

# Photon Counting & the Statistics of Light

September 22, 2014

Jung Lin (Doris) Lee  
dorislee@berkeley.edu

Group partners: Jennifer Ito, Manuel Silvia  
Prof. James Graham, UGSI Heechan Yuk, Isaac Domagalski

---

## Abstract

We present the methods used to conduct the photon counting experiment and the statistical results obtained from the collected data. By varying the brightness of the light source, we found that photon arrival rate increases, as seen from the decreased number of events. (modify this??) Analysis on a single data set shows that as the the number of events in an experiment increases, the sample mean( $\bar{x}$ ) converges to the theoretical mean( $\mu$ ). However since systematic effects can not be eliminated by increasing number of events, we mitigate the effect of afterpulsing by vetoing photon arrival events that occurred immediately after another. The processed data results in a histogram that align very closely to exponential distribution. In addition, binned data for a fixed time interval also matches closely with the Poisson distribution. Therefore, by fitting the distribution we obtain the sample mean parameter of the experiment.

---

## 1. Introduction

Photon statistics are the fundamental basis to optical astronomy since they constitute the data we collect in optical astronomy. are important to astronomy because optical astronomy can be used to study galaxy morphology, individuals sources or conduct statistical studies for cosmological studies. Large sky surveys such as the Sloan Digital Sky Survey (SDSS; York et al.) and the Dark Energy Survey (DES; The Dark Energy Collaboration) imaging data in the optical wavelength.

In this experiment, we recorded the arrival of photons from an LED and analyze In section 2, I will describe the experimental method that my group used to take the data and the results that we obtained. Subsequently, section 3 introduces the various statistical measures that we can use to describe the findings and observation on a single dataset, including the effect of binning the data into histograms. Section 4 describes the two prob-

ability density functions that are relevant to these histograms and dataset.

## 2. Experiments

### 2.1. Methods

The arrival of a photon is registered by the photomultiplier tube (PMT), and data was acquired using the CoinPro software. The resulting processed data file is a series of time stamps of when the data arrived. Since CoinPro is ran on a 32-bit machine, we can see the cycling through of time in FIG REF. Although absolute time stamp is the recorded quantity, it is often more useful to consider the time elapsed between two consecutive events ( $\Delta t$ ) as plotted in Fig. in order to get a sense of the photon arrival rate for some sample waiting time.

Sometimes noise pulses resembling Fig.2 can be observed after the original pulse. These events are caused by the ionization of residual gas molecules

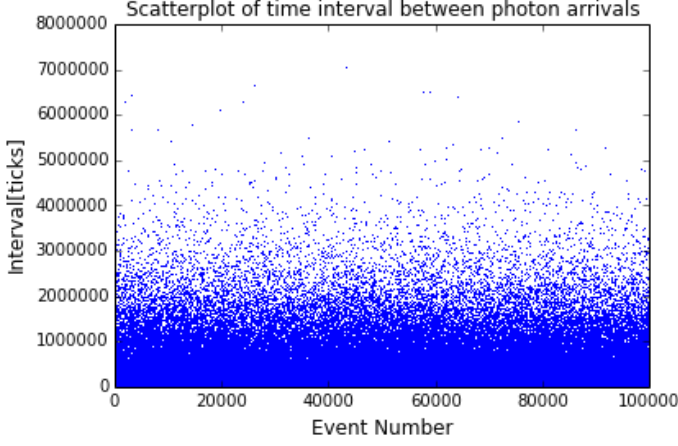


Figure 1: With the removal of after pulse event, we can see that there is a decline in the number of events. This makes sense because we are doing an interval cut at 3000 clock ticks, which should result in a clear boundary at  $y=3000$  but this can not be seen due to the large range of interval value that the y-axis spans over. The data representation in Fig.3 and 6 better shows the effect of gating afterpulses.

in the PMT's vacuum chamber. We eliminate this systematic effects by cutting away data with subsequent pulses lie within the time frame of  $2.5\mu s$  after an original pulse as shown in Fig.6.

## 2.2. Changing LED Brightness

By listening to ticks on the squawker, there was no detectable difference from the 3 o'clock off position to the 12 o'clock position. Therefore, we started off at the 12 o'clock location and went to the max location ( $\approx 90^\circ$ ), diving this region into 6 equal parts <sup>1</sup> is roughly  $18^\circ$  increments. Even though the measurement of the LED brightness is eyeballed and not that precise, it did not significantly affect the plot since the the data point should fall on the  $\sigma = \bar{x}$  line since the  $\sigma$  and  $\mu$  of an exponential distribution is always equal.

<sup>1</sup>Although it seems like only 5 data points are shown in the plot, this is merely an artifact of the choice of data representation as circles as the first two data points lie so closely that they are almost indistinguishable. This is easily verified by using the dot representation with matplotlib's 'setting

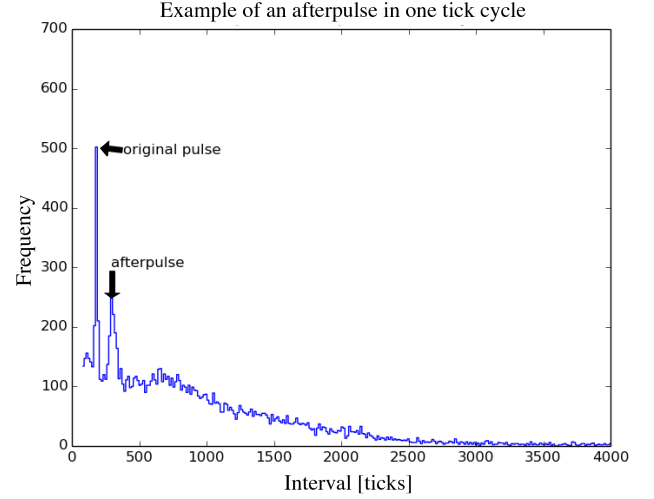


Figure 2: With a histogram of bin size = 500,000, the two peaks (original and afterpulse) as well as a gradual decline of ticks in the next 3000 clock ticks can be seen here.

Intensity[°]	Final time stamp in time series [ticks]
0	86447071309
18	68475846523
36	35127973096
54	7517634481
72	5787954870
90 (max)	5558407093

Table 1: Intensity is specified as the LED control knob position in degrees, as described earlier in Sec.2.2. Since we are taking a fixed number of photon arrival events, as the brightness increases, the faster photon arrival rate means that 10,000 event count finishes faster than it did for lower brightness; thus, accounting for the smaller time stamp values.

## 3. Descriptive Statistics and Histograms

Using the  $N=100,000$  dataset with LED on the maximum brightness setting, we obtain a sample mean of 561571.83 and the standard deviation is 637988.76 on the whole dataset of 99833 events <sup>2</sup>. This is the most precise measure of the statistical quantity for the data collected since the more data that we consider when computing these quantitative, the closer the sample mean converges to the mean value as shown in Fig. 5. Higher precision measurement of  $\bar{x}$  and  $\sigma$  can be obtained by tak-

<sup>2</sup>CoinPro does not stop the event count at exactly 100,000 events.

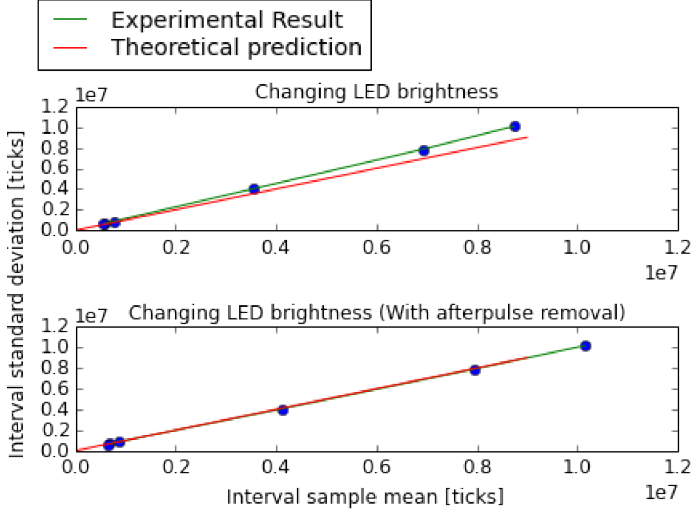


Figure 3: The slope of the line should be one since both the standard deviation and the sample mean of each dataset is equal to the single-parameter  $\tau$  in the exponential distribution. We find that gating the afterpulses causes the  $\bar{x}$  to get closer to  $\mu$ .

ing a sample of larger  $N$  so that  $\bar{x} \rightarrow \mu$ , but this does not ameliorate any systematic effect such as dark current or afterpulses.

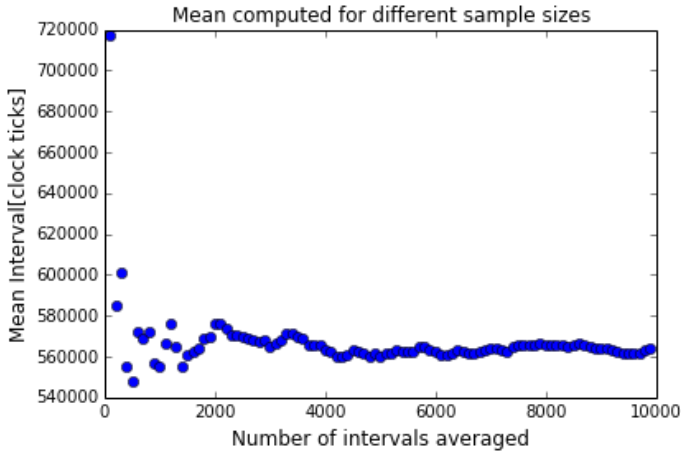


Figure 4: Large fluctuation in computed mean decreases as we use larger datasets to compute the mean value.

### 3.1. Standard Deviation of the Mean (SDOM)

## 4. Photon Statistics

### 4.1. Exponential Distribution

Exponential distribution described by Eq.1 governs the histogram of  $d\mathbf{t}$ , the elapsed time between photon arrival event.

$$P(t, \tau) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \quad (1)$$

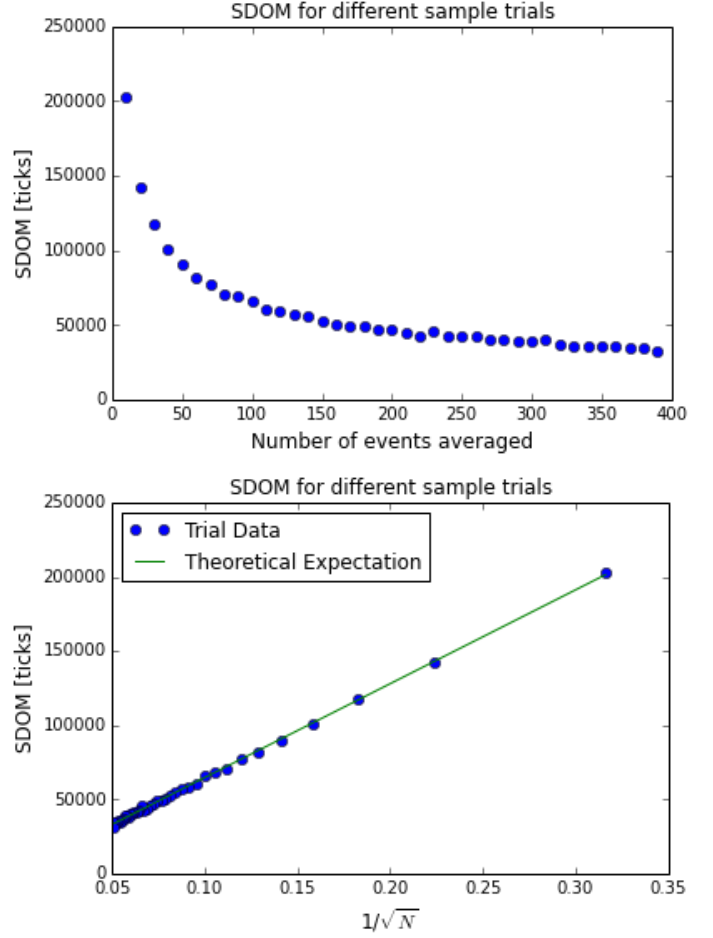


Figure 5: SDOM

A defining characteristic of the exponential distribution is that it is continuous “memorylessness”. In this case, the photon arrival rate at some moment in time,  $t$ , is not dependent on the photon arrival rate prior to time  $t$ . The green curve in Fig.6 is plotted from the analytical solution shown in Eq.2, which is obtained from integrating the exponential distribution with a step-size  $\Delta t = 1.42 \times 10^7$  and  $\tau$  set as the mean of the  $d\mathbf{t}$  dataset.

$$\int_{\tau}^{t+\Delta t} \frac{1}{\tau} e^{-\frac{t}{\tau}} dt = -e^{-\frac{t+\Delta t}{\tau}} + e^{-\frac{t}{\tau}} \quad (2)$$

Since the exponential probability density function (PDF) is normalized to one, this result is scaled by the number of events to fit the histogram.

### 4.2. Poisson Distribution

Although closely-related, the Poisson distribution is unlike the exponential distribution in that

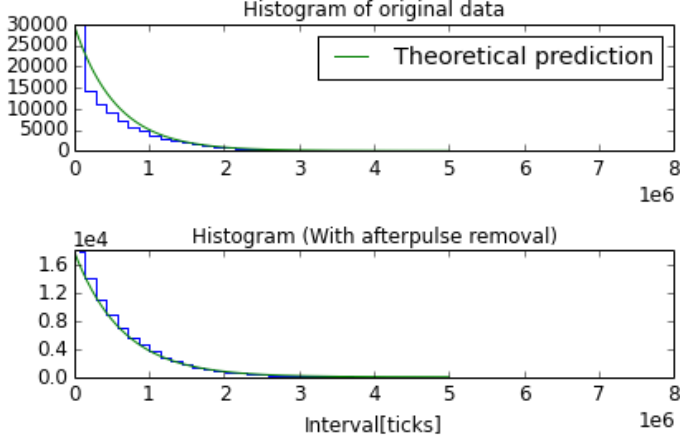


Figure 6: Histogram of bin size = 50 before and after gating afterpulses, with the scaled exponential PDF fit.

it is a discrete probability distribution that arises from measuring the number of events that occur within a fixed sampling interval. Both underlying process applies to rare event for very large  $N$ , but the experiment method is different. In this case, the experiment is the same, but we change the way the data is represented, as binned data for a fixed time interval.

## 5. Conclusion

Our findings can be summarized to the following points:

- 
- Since photon arrival is rare and random in the limit of large  $N$ , it can be closely follow a Poisson distribution.
- 

In this experiment, we only attempt to mitigate the effect of one type of systematic error vetoing the afterpulses. One possible extensions to this project may be to investigate the effect of other types of systematic error such as dark counts and thermal noise on the quality of the data. This effect can be quantified by measuring the signal to noise ratio and comparisons among different types of PMT.

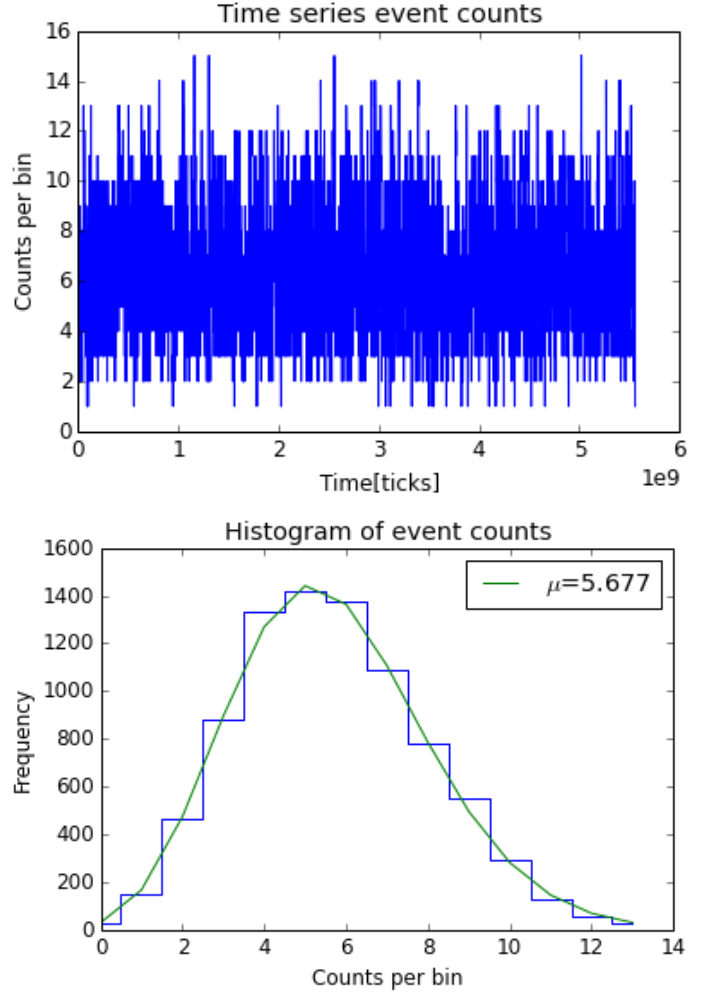


Figure 7: The bin counts are obtained by counting the number of photon arrival events that lie within an interval of 3.09 milliseconds. From the time series plot, we see that most time intervals counted 4 to 8 events. This is confirmed by using a Poisson curve fit on the histogram where the value of the fitting parameter is 5.677.